Image Processing for High-Precision Eye-Movement Tracking

Jeffrey B. Mulligan

Eye movements can provide a wealth of information about how human operators perceive and process visual information. Video-based measurement systems offer considerable advantages over competing approaches for use in applied contexts outside the laboratory: first, they do not require physical contact with the eye, and second, they do not restrict the subject's movement. This research investigates image processing methods in an effort to obtain greater accuracy from video images of the eye.

Currently available commercial systems use special hardware to compute eye position in real time from images of the eye's anterior structures (first figure). To attain real-time performance, these systems must use relatively simple processing algorithms. This project is concerned with maximizing the final accuracy by the application of more sophisticated image processing, even when the calculations cannot be performed in real time on present-day microcomputers.

The first challenge was to develop efficient and convenient methods for continuous digital recording of video data. Two approaches to this problem have shown promise. In one, the application of hardware image compression reduces the data rate to one easily handled by a single computer disk. Results obtained using simulated data indicate that moderate levels of compression have negligible effects on the final accuracy. In the other, an array of parallel disk drives achieves a data rate capable of storing video with no compression.

Digital storage of the video images and off-line processing allow sophisticated processing algorithms to be applied. A particularly difficult problem has been reliable tracking of the fourth Purkinje image, formed by reflection of the illuminator from the posterior surface of the crystalline lens of the eye (the small bright spot in the first figure). The position of the fourth Purkinje image is especially useful, because the distance between this feature and the bright corneal reflex, or first Purkinje image (the large bright spot in the figure), is directly related to the direction of gaze, unlike the positions of the individual features that confound translations of the head with shifts of gaze directions. Difficulties in tracking

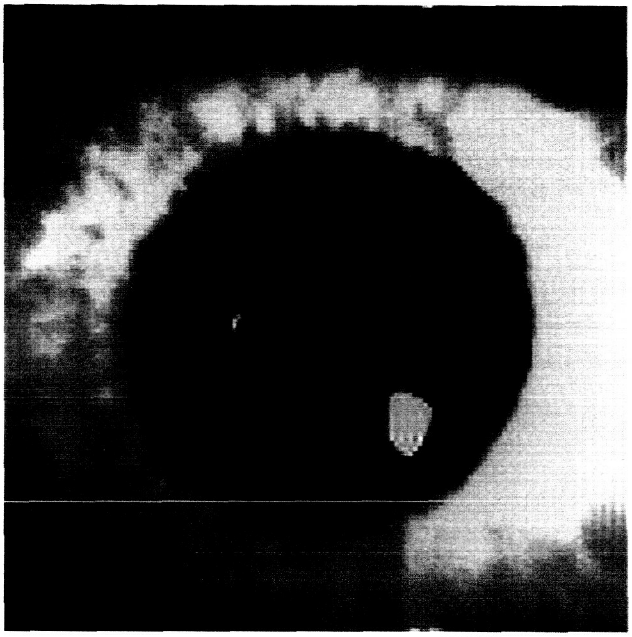

Fig. 1. Subsampled video image of the anterior structures of the eye. The central dark disk is the pupil, formed by the edge of the iris. The two bright spots within the pupil are the first (large) and fourth (small) Purkinje images, whose relative positions give a precise indication of gaze direction.

the fourth Purkinje image arise from its small size and low contrast. This project has now succeeded in tracking the fourth Purkinje image by using a Gaussian curvature computation on an image from which the corneal reflex and pupil margin have been masked off.

Although images of the pupil are easy to obtain simply by pointing a camera at the eye, greater accuracy still can be achieved by ophthalmoscopic imaging of the retina (second figure). This requires a special optical setup, and currently requires that the subject's head be stationary; however, it permits much greater optical magnification, with consequent increases in resolution and accuracy for small eye movements. This project has developed a computer program that constructs a large retinal mosaic for each subject. Using techniques similar to those used

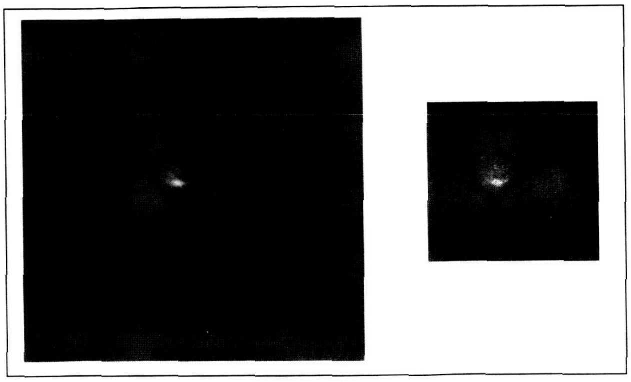

Fig. 2. Retinal images obtained with a table-top video ophthalmoscope. The left-hand panel shows a subsampled version of a single video field, in which a faint image of the optic disk and retinal blood vessels may be seen on a large background of camera noise. The right-hand panel shows a composite image constructed by registering and averaging approximately 1000 images like the one shown in the left panel.

in the analysis of satellite imagery, the program registers and averages a large number of images, each of which covers just a small part of the subject's retina. Once this mosaic or template has been constructed, subsequent records are analyzed by registering the individual frames to the template. In addition to images obtained in the laboratory, the software has been applied to images obtained with a scanning laser ophthalmoscope, a clinical instrument used in the diagnosis of various ocular disorders.

For both classes of imagery described, the software developed in this project provide a level of accuracy commensurate with the inherent physiological noise (about 1 arc minute), significantly better than commercially available video-based systems and comparable to the best invasive methods.

Point of Contact: J. Mulligan
(650) 604-3745
jmulligan@mail.arc.nasa.gov

Measuring Human Detection Templates

Albert J. Ahumada, Jr., Andrew B. Watson, Bettina L. Beard

As part of NASA's goal to improve aircraft safety and performance, Ames is developing models that can be used to predict a human observer's ability to detect visual targets. Often these models perform well because they mimic visual system processing. When task performance depends on the observer's memory of a target, these models should include a characterization of these internal representations, or memory templates.

A technique was developed to measure these templates when an observer is discriminating two different targets. The technique involves adding a small amount of random noise to each target stimulus. The two targets are then presented to the observer. The observer's discrimination response is then correlated with the lightness or darkness of the noise at each location, or pixel, in the image. If a particular pixel significantly contributes to the observer's decision, then the resulting response correlation image will show lighter or darker areas in those regions. The response correlation image illustrates the contribution of each noise image pixel to the observer's decision and can represent which features of the stimulus are being used to make the discrimination.

At Ames, this technique was demonstrated using a well-studied visual discrimination task, vernier acuity. Vernier acuity refers to the smallest misalignment of two lines that an observer can detect. The first figure illustrates the two stimuli presented in a vernier acuity discrimination task. One stimulus is a pair of lines in alignment. The other is the same except the right line is elevated by a single pixel (0.005 degree of visual angle). Human observers are quite precise in detecting misalignment and it is of interest to know what features of the stimulus are used that can account for such precise discriminations. This knowledge can greatly improve model predictions of target detection and discrimination.

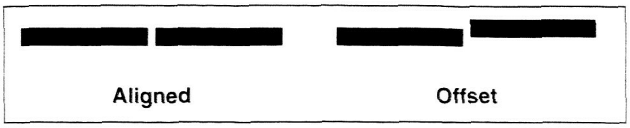

Fig. 1. Vernier acuity stimulus. In this example, the vernier stimulus is composed of two line features that are 0.02 degree of visual angle in length. On each trial, the right line will either be aligned with the left, or displaced upward relative to the left line. The task is to categorize the trial as "aligned" or "offset."

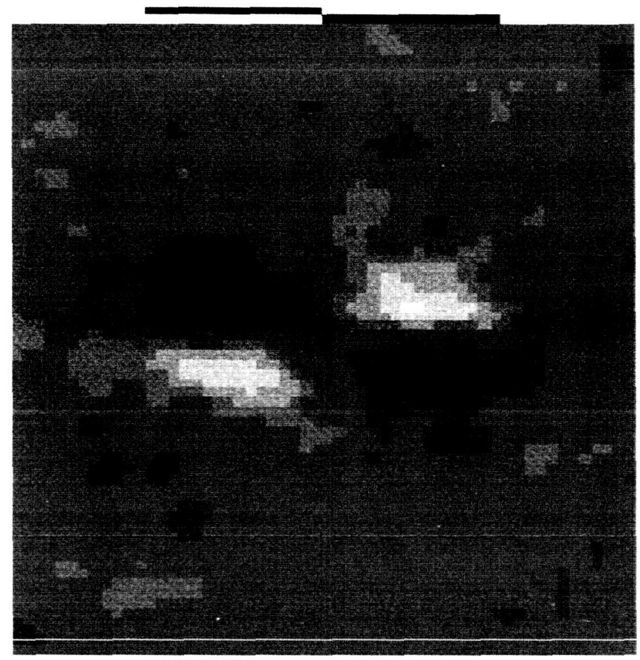

Fig. 2. Response correlation images for a vernier acuity task. A response correlation image is shown for the combined data sets of three observers. Dark areas mean that darker noise pixels in these locations led to more "offset" responses. Light areas in the image mean that lighter pixels led to the "offset" response.

The second figure shows the response correlation image obtained after correlating the added noise pixel values with the observer responses. Current visual discrimination models that mimic visual system processing, but ignore observer templates, typically predict that the right side of the image should show a blurred version of the difference stimulus, as does appear. However, these models predict that nothing should appear around the left vernier feature since the images are the same in this region. The response correlation image shows that contrary to this prediction, the observers pay approximately equal attention to the fixed line on the left.

The response correlation technique can make an important contribution to current visual discrimination models by improving predictions of target detectability. This technique is useful for a variety of tasks in clarifying the underlying features used to form and upgrade memory templates.

Point of Contact: A. Ahumada, Jr.
(650) 604-6257
aahumada@mail.arc.nasa.gov